BBC RD 1985/4

RESEARCH DEPARTMENT REPORT

Electronic detection and concealment of film dirt

R. Storey, B.Sc.

Research Department, Engineering Division
THE BRITISH BROADCASTING CORPORATION February 1985

BBC RD 1985/4
UDC 621.397.3.004.64

ELECTRONIC DETECTION AND CONCEALMENT OF FILM DIRT

R. Storey B.Sc., C.Eng., M.I.E.E.

Summary

Film as an image source for television has four obvious impairments; film grain, dirt, unsteadiness and motion aliasing. The first three of these impairments can, at least in theory, be minimised electronically. For example, a previous Report has described the design of specialised electronic processing equipment to reduce the level of noise in a television signal; this is also very effective at reducing the visibility of film grain.

One method for detecting the presence of dirt on film passing through a telecine, using infra-red light, is described in another Report. However, it requires a specialised telecine and can only detect dirt particles adhering to colour film stock. A second, totally new method, which forms the subject of this Report, is entirely electronic and operates directly on the coded video signal from a standard telecine. It can detect and conceal the image of dirt particles and printed dirt on both monochrome and colour film.

The methods of detection and concealment are described together with details of a preferred realisation of the equipment. A description of the unsuccessful techniques examined during the course of this investigation is also included, for completeness.

The prototype equipment has a single adjustable parameter which governs the maximum detectable size of dirt for a given visibility of motion impairment. The small dirt only mode may be safely used without previewing, when it will conceal the majority of dirt in typical 16 mm film with negligible motion impairment. The remaining large dirt can be concealed by adjusting the motion protection compromise manually, but source material should then be previewed, to check for motion of the type likely to cause noticeable motion impairment.

Issued under the Authority of

M. E. B. Moffat

Head of Research Department

Research Department, Engineering Division
BRITISH BROADCASTING CORPORATION
February 1985

(PH-268)

It uses SI units in accordance with B.S. document PD 5686.

ELECTRONIC DETECTION AND CONCEALMENT OF FILM DIRT

R. Storey B.Sc., C.Eng., M.I.E.E.

ELECTRONIC DETECTION AND CONCEALMENT OF FILM DIRT

R. Storey B.Sc., C.Eng., M.I.E.E.

Introduction

Programmes from film form a substantial proportion of the output of both B.B.C. TV networks, and are likely to continue to do so for the foreseeable future. Technological progress is causing a steady shift towards electronically generated images but film will remain a major source, if only in the shape of archive and feature film material.

There are four obvious impairments which are frequently seen on film programmes: these are, film grain, dirt, unsteadiness and motion aliasing. The first three of these impairments are steadily being reduced by improvements in film stock, handling techniques and camera, printing and telecine hardware. The effects of these three impairments can, however, be reduced still further, by a number of processing techniques.

The visibility of film grain has been substantially reduced by post processing with a video noise reducer[1]. Recently, film dirt has been detected and concealed by both an optical and an electronic method; the latter forms the subject of this report. Unsteadiness remains to be investigated and motion aliasing is a fundamental and probably incurable, failing caused by the low frame rate used for film.

Two methods of dirt detection have been investigated, the first being optical and requiring a specially designed telecine[2]. This technique relies on the transparency of colour film emulsions to infra-red light, allowing dirt particles adhering to the film to be detected by an infra-red sensor. It cannot detect dirt on monochrome film stock because the silver image is opaque to infrared light, nor can it detect the printed image of dirt on negative film stock. Infra-red dirt detection can be an integral part of the new generation of solid state line array telecines, but practical problems prevent its incorporation into existing machines. Furthermore, limitations of lens performance at infra-red make optical dirt detection less effective for small dirt than for large dirt.

As an alternative that can be used with existing telecines, an electronic method of film dirt detection and concealment has been developed and forms the subject of this Report, and an associated Patent Application[3]. The output of the telecine is inspected electronically for signal components having the characteristics of film dirt. The detection scheme draws no distinction between printed dirt and dirt particles adhering to the film in the telecine, being equally effective for both. There is, however, an adjustable compromise between dirt detection effectiveness and motion impairments. The compromise ranges between detection of small dirt only with negligible motion impairment, to detection of all sizes of dirt with occasional motion impairment, and must be set by an operator for the best subjective effect. It should be noted that in its small dirt mode, the electronic detection system complements the infra-red technique, which is best suited to finding large dirt. The two systems combined, can therefore conceal dirt of any size with negligible motion impairments.

The equipment has been designed to operate on composite PAL signals making it suitable for use as a central facility in a telecine area. It has a propagation delay of three TV fields which in an operational environment, probably limits its use to telecine-to-videotape or videotape-to-videotape transfer.

1. Electronic film dirt detection strategy

The first stage in solving a problem of this kind is to assess its nature in detail, listing all the known characteristics of film dirt and the pictures in which it occurs. From this information, the most promising strategy can be chosen. The evolution of this strategy assumes that the film images have been converted into a television signal in a conventional telecine machine.

Consider an isolated dirt particle adhering to a film frame, which is one of a sequence describing a moving scene. The transmission of light through the film is reduced in the area obscured by the dirt particle, causing a dark spot to appear momentarily, when the film frame is displayed. Conversely a dirt particle adhering to a negative film stock would appear as a bright spot when printed or reproduced as a positive image.

In the general case, a scene may contain motion and unsteadiness. It will also suffer from varying amounts of film grain[1]. No direct access is possible to the signal value that would exist in the absence of

dirt nor can anything categorical be said about the size or shape of dirt or the amount of signal modification it causes.

The human visual system is finely tuned to the recognition of anything transistory or out of place, no matter how complex the remainder of the scene. For this reason, film dirt is immediately obvious to an observer, even though each particle appears for only 40 milliseconds.

An observer understands the content of a scene and the permissible motions of objects in it. This process of motion understanding involves more than image correlation, since an object can change its shape totally while in motion (e.g. a bird's wing). It requires in addition, a knowledge and recognition of all the possible aspects of a moving object. Clearly any attempt to predict the true scene content by modelling the processes used in the human visual system, will be doomed to failure because of their complexity.

If a practical strategy for detecting film dirt electronically is to be developed, an engineering compromise is required. A starting point can be found by considering the simplest possible case, a scene containing no motion or unsteadiness and a known level of film grain. Under these conditions, a fairly accurate estimate of the true signal value in any film frame can be derived by interpolation from preceding and succeeding reference frames. Dirt can now be detected electronically, by looking for signal values in a frame which differ from the predicted values, by an amount greater than the uncertainty caused by grain in the reference frames and in the frame under inspection. (See Fig. 1).

This basic detection algorithm is termed a "Dirt Signature Detector" and is capable of finding almost all the dirt in a stationary scene, failing only in the unlikely event of successive frames containing spatially coincident dirt. Its effectiveness for dirt detection in realistic pictures, however, depends upon the accuracy of prediction in the presence of motion, unsteadiness and other artefacts found in typical film pictures. Its tendency is to misinterpret these artefacts as dirt.

The ideal way to generate an accurate prediction would be to model the mental processes of a human observer to create an expected scene from information in the surrounding frames. This is theoretically feasible for simple uniform global motion, although the processing power required makes even this impracticable. Realistic pictures, however, contain motion of foreground against background, object rotation, and a host of similar transformations that are not yet open to real-time analysis in the electronic signal domain.

The problem therefore, reduces to one of making a prediction which is accurate for most of the time, and to prevent it from failing noticeably in the presence of complex motion and unsteadiness. This is partially achieved in practice, by using the algorithm of Fig. 1, but confined to the immediately preceding and succeeding reference frames (See Fig. 2), to minimise the persistence of any prediction errors. The threshold parameter δ is used as a "trend removal" element to effectively modify the expected prediction *error* as a function of both spatial and temporal scene content. This approach guarantees that the "dirt signature algorithm" imposes no limits on the size or shape of dirt that can be detected; the only constraints appear in the formulation of a suitable trend removal signal, and this can be tailored to suit the application.

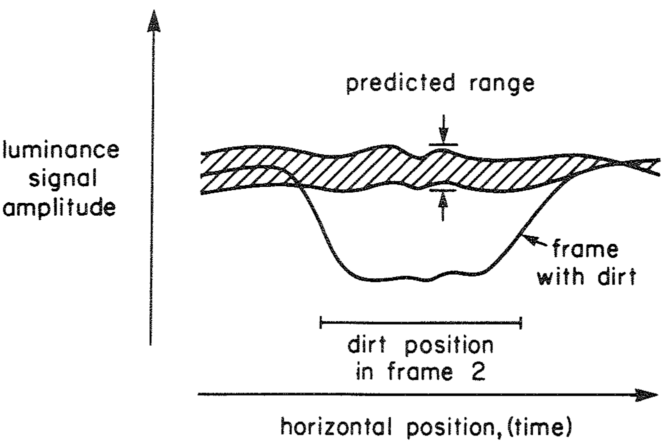
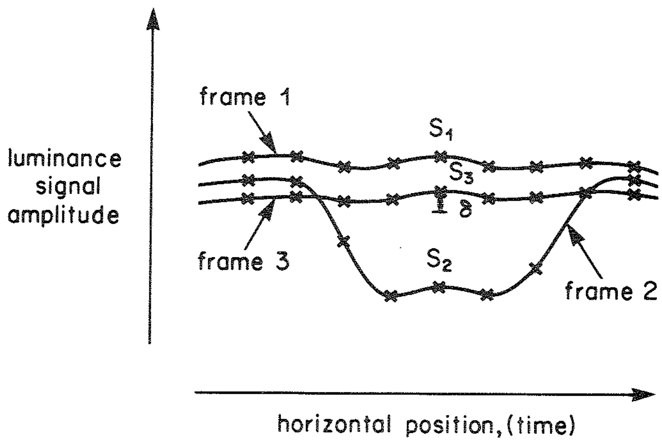


Fig. 1 – Variation of luminance from the predicted range caused by scanning through a piece of dirt.



$$\left.\begin{array}{l} |S_3 - S_2| > \delta \\ |S_1 - S_2| > \delta \\ \mathrm{sgn}(S_3 - S_2) = \mathrm{sgn}(S_1 - S_2) \end{array}\right\} \begin{array}{l}\text{dirt}\\ \text{signature}\\ \text{algorithm}\end{array}$$

Fig. 2 – Coincident horizontal scans in three consecutive frames with dirt in frame 2 only.

2. Hardware vs software

Past experience[1] has shown that picture processing schemes must be tested on a wide variety of source material. The range of scene stimuli in typical pictures is so enormous, that an assessment over a limited range of material may easily show a processing scheme to be foolproof, only to reveal catastrophic failures when the equipment is subsequently used on a wider range of material. This, coupled with the prime requirement for a film dirt detection scheme that it should not fail noticeably for any motion or unsteadiness stimulus, dictates that assessment should occur in "real-time" and over a wide range of source material.

Assessment of continuous moving sequences implies that all signal processing must be done in "real-time" also. Current practical limitations force the use of dedicated hardware for signal processing rather than more flexible software techniques. The main disadvantages of a dedicated hardware approach, is that each new idea or technique requires construction of new hardware. If however, the number of unsuccessful attempts can be minimised, the equipment developed in the process of finding a satisfactory solution can form a field trial prototype.

3. Evolution of the "trend removal" algorithms

Section 1 describes an algorithm which looks for the temporal signature of film dirt. This algorithm constitutes a "Dirt Signature Detector" which is capable of finding practically all the visible dirt in stationary pictures. It does not misinterpret shot changes for dirt nor, with the aid of a fixed threshold parameter, is it confused by film grain. It does, however, misinterpret the motion of some types of detail, since moving detail will inevitably cause the dirt signature conditions to be satisfied along the temporal axis of some pixels.

It is not generally feasible to predict the exact value of a pixel in the presence of moving detail using currently available techniques. The best that can be achieved is a value for the expected maximum prediction error caused by motion at each pixel, estimated from surrounding pixels in both space and time. The estimate should ideally represent the motion trend surrounding a pixel.

The necessary elements of a successful trend removal function were found by observing the performance of the basic dirt signature detector, used with a fixed threshold, on a wide range of source material. The resulting trend removal scheme allowed the predominant sources of failure to be reduced to an acceptable level, and revealed in the process, some second order effects hitherto concealed.

The primary sources of failure were motion of fine detail caused by film unsteadiness, and general motion of objects within the scene. The trend removal signal components, one being a function of film unsteadiness and the other a function of motion, are described separately.

3.1. Unsteadiness

Unsteadiness in film pictures is caused by misregistration of successive images, either in the camera, during printing, in subsequent conversion to electrical signals or a combination of all three. It shows itself as a bodily displacement of the image in random directions by random but usually very small distances.

Failure of the "Dirt Signature" algorithm manifests itself as misinterpretation of unsteady fine detail for dirt. This appears in the concealed picture as a coarse grained, erratic change in the texture of detail.

The problem is to derive an electrical signal, u, describing the amount of unsteadiness between the frame under inspection and one of the reference frames. This unsteadiness signal can be used to increase the amplitude of expected error in detailed areas when there is unsteadiness between the frame under inspection and an adjacent reference frame.

3.1.1. The unsteadiness quantifier

A measure of the amount of motion between two film frames can be derived by integrating the modulus of the difference between the television signals describing the two frames, over a complete frame scan.

This arrangement is shown in Fig. 3. A picture delay provides access to the signal "A" describing the luminance at a point in the current frame and a signal "B" describing the same point in the previous frame. Any difference caused by relative motion between the two frames will cause a finite output from the subtractor. All such outputs are made positive by the rectifier and integrated through a full frame period. Provided that the motion was less than the smallest resolveable picture element (termed "a pixel") as is normally the case, the integral value will increase with increasing displacement. If, however, the displacement is greater than a pixel, the relationship between integral value and

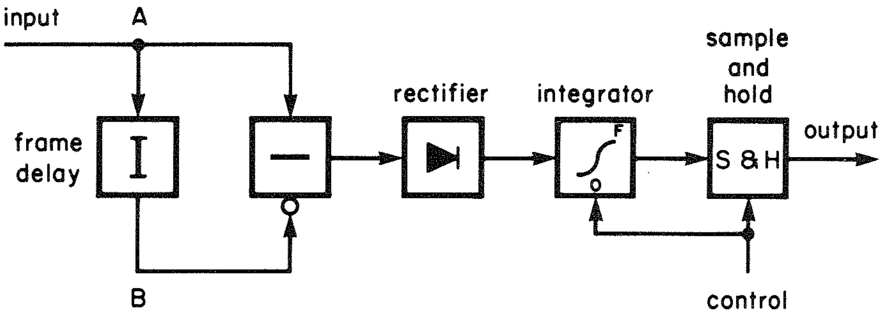


Fig. 3 – Simple motion quantifier.

displacement need not be monotonic but will generally be non-zero. Fortunately, the image displacements found in television film pictures are generally very small, typically a fraction of a television line pitch vertically or a similar distance in any other direction.

The integral value in Fig. 3 which is sampled at the end of a frame period, is therefore monotonically related to the unsteadiness displacement between the two input film frames. It is also proportional to the amount of high frequency detail in the film frames. This occurs because a displaced plain area generates a vanishingly small rectified frame difference, whereas a displaced finely detailed area generates a large rectified frame difference. A normalising operation must therefore be applied to the rectified frame differences before they are integrated. An unsteadiness quantifier, normalised for fine detail, is shown in Fig. 4.

Two detail detectors measure the amount of detail surrounding the pixel values appearing at the two subtractor inputs. The two detail signals are added and the sum is applied to a divider whose dividend is the rectified frame difference.

For small displacements, the resulting quotient is monotonically related to displacement and independent of detail, as is the unsteadiness value "u", derived by integrating the quotient over a frame period and sampling that integral at the end of the frame.

3.1.2. The detail detector

The detail detector is required to measure the amount of signal variation within one pixel of the point from which a frame difference is taken.

This is achieved using the system shown in Fig. 5. There are four cascaded delays, two of length T, the time taken to scan a pixel, and two of length H – T where H is the horizontal scan period. These delays give access to signal values immediately above and below the centre value S_{00}, and a pixel on either side. The detail signal is derived by summing the moduli of differences between S_{00} and the surrounding points.

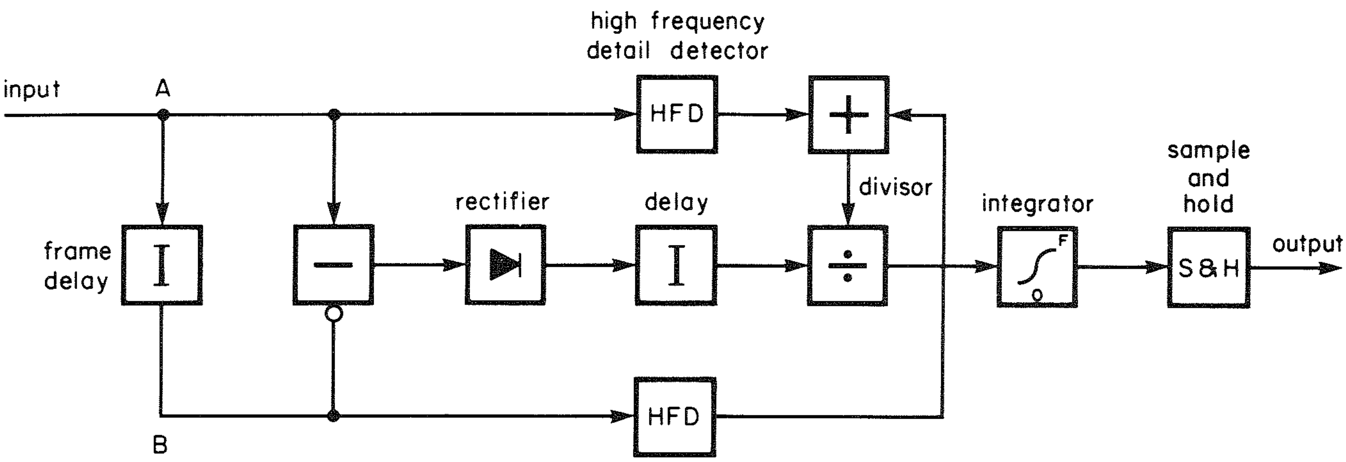
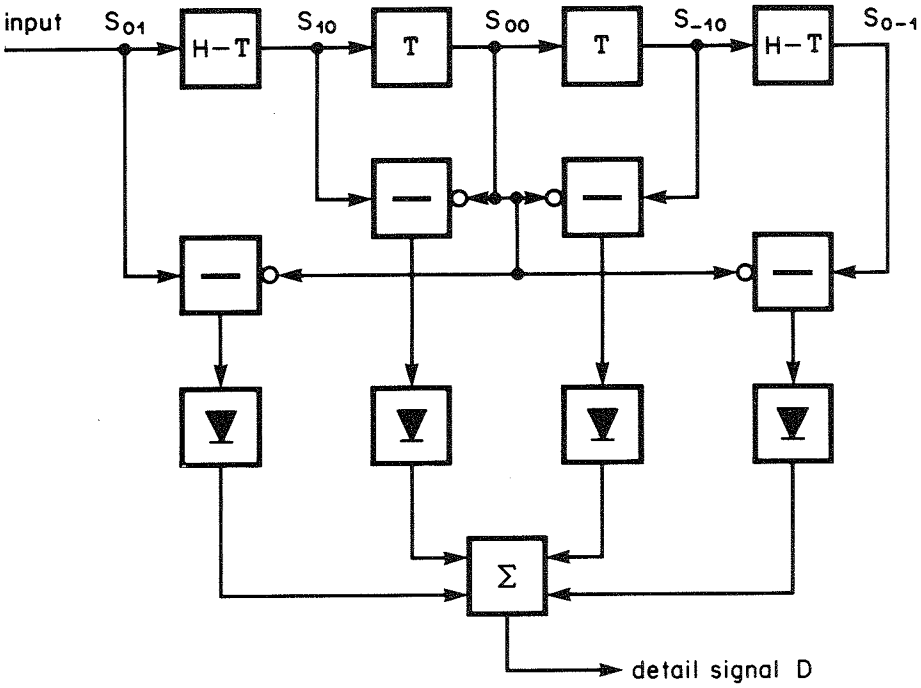


Fig. 4 – Unsteadiness quantifier with detail normalisation.



H = 1 line delay
T = 1 element delay
$D = \sum |S_{10} - S_{00}| + |S_{-10} - S_{00}| + |S_{01} - S_{00}| + |S_{0-1} - S_{00}|$

Fig. 5 – Detail detector.

3.2. Motion protection

Generation of a trend removal component to prevent impairments caused by misinterpretation of general motion in a scene, is a more difficult problem to solve than that of the unsteadiness component. This is because the range of motion displacements commonly found between consecutive film frames is large. Also the sampled nature of film pictures and scene integration caused by a finite camera shutter time become significant for large displacements.

Five motion detection schemes were tried during the development, each one having its strengths and weaknesses. Two versions of a classical motion detector[1,4] were finally chosen and combined adaptively to give a system capable of protecting against the majority of scene motion.

The following description is of the final motion detector followed, for completeness, by a description of the other schemes investigated.

3.2.1. Classical motion detector

Motion of scene content between two frames of a film sequence, can be detected by subtracting the signals representing corresponding picture elements (pixels) in the two frames from one another. The difference in pixel values in which no motion has occurred is generally small, being caused primarily by film grain. Conversely, pixels representing moving scene content or newly revealed background generally produce larger differences.

The exact nature of an inter-frame difference signal is impossible to describe analytically because scene content and motion are not generally predictable. It is possible however, to derive useful information from an inter-frame difference signal by using well known techniques for non-coherent signals, based on power measurement. The behaviour of subsequent processing of an inter-frame difference signal can also be estimated by assuming an idealised stimulus, such as a moving plain object against a stationary background, and verified experimentally for real moving images.

A movement detector which effectively measures the signal power in the inter-frame difference signal caused by motion is shown in Fig. 6. It is instructive to consider the behaviour of this circuit for the idealised case of a circular object moving against a plain background, where the difference signal is taken from two consecutive film frames. Figure 7(a) shows the image recorded on each film frame; note that, since the film was exposed to the scene for approximately half the time separating the two frames, each recorded image is elongated in the direction of motion. Any fine detail on the moving object would be blurred by the motion; this effect is termed camera integration and occurs because the luminance recorded at each point on the film emulsion is proportional to the integral of the light falling on that point while the camera shutter is open.

Figure 7(b) shows the signals corresponding to a horizontal scan through the image of the moving object and Fig. 7(c) shows the inter-frame difference

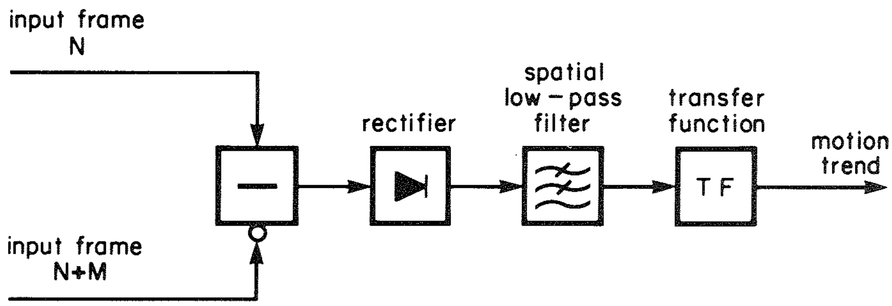


Fig. 6 – Classical motion detector.

signal. The difference signal can be positive or negative and, for a plain moving object, has a reduced amplitude in the area of overlap. If the frame difference signal is rectified, Fig. 7(d) and passed through a low pass spatial filter, Fig. 7(e), the resulting waveform is always positive and encompasses both image positions; satisfying the primary requirements for a movement trend signal.

3.2.2. Reference frame motion detection

Consider now, three consecutive film frames captured from a moving sequence as in Fig. 8. A

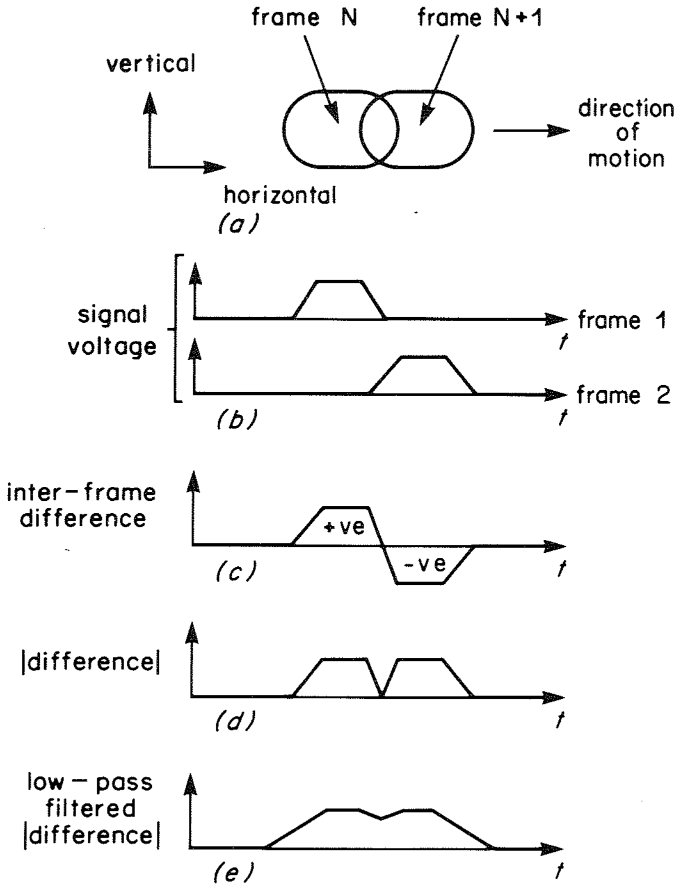


Fig. 7 – Idealised waveforms caused by a moving object.

movement trend signal is required to prevent misinterpretation of moving scene content in the centre frame, while it is being examined for dirt by the dirt signature algorithm of Fig. 2. If the centre frame (frame 2) signal is used as an input to the motion detector, pieces of dirt whose dimensions are comparable to or greater than the spatial filter aperture will contribute to the motion trend signal and prevent their own detection. Detection of all sizes of dirt can be preserved by using only the two adjacent "reference" frames (frames 1 & 3) as inputs to the motion detector. This is termed a "Reference frame motion detector" and produces a very effective motion trend signal in the absence of rapid or erratic motion.

If, however, the scene contains objects moving fast enough for their images on successive frames to be separate as in Fig. 9(a), the reference frame motion detector output, Fig. 9(b) does not cover the position of the object in Frame 2 and, since this image has the same signature as dirt, it is misinterpreted. This failing can be reduced to some extent by enlarging the spatial filter aperture and thereby spreading the output response as in figure 9(c).

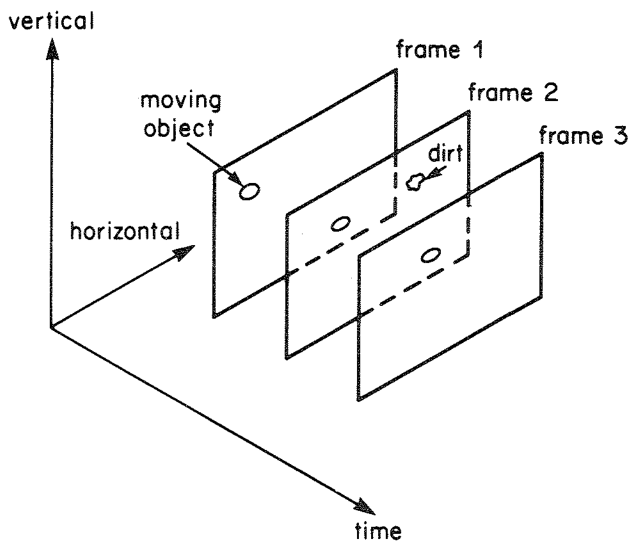


Fig. 8 – Three consecutive film frames.

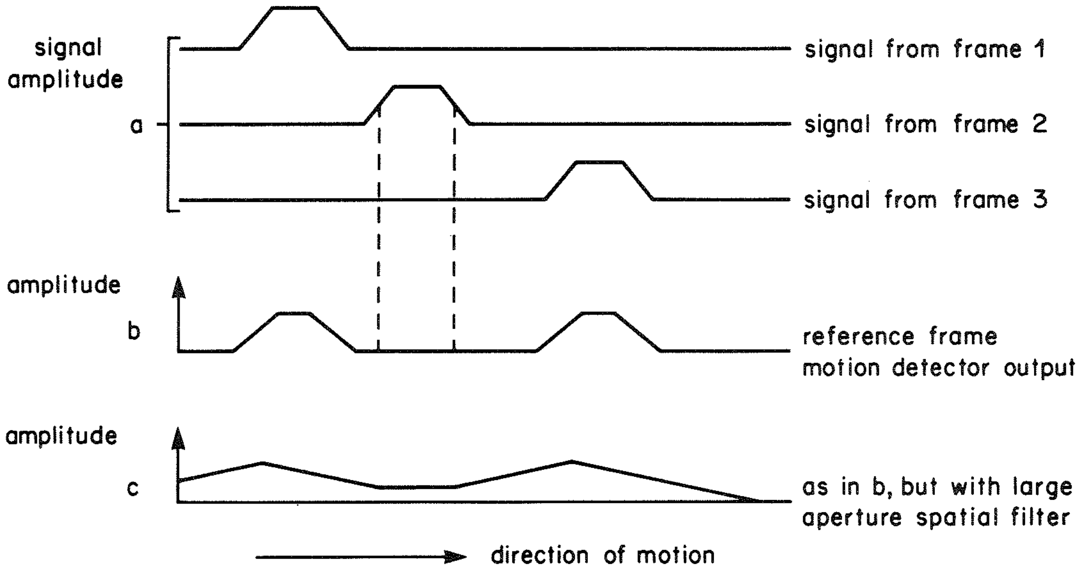


Fig. 9 – Idealised "reference frame motion detector" response to rapid motion.

However, the amount of aperture enlargement required to be effective with typical fast motion is so large, that an unacceptable halo of undetected dirt appears around the edges of moving objects. A reference frame motion detector can thus be used only for scenes containing slow motion but it does allow detection of all sizes of dirt.

3.2.3. Centre frame motion detection

We have seen in the previous section, that protection of rapid motion requires a contribution to the motion detector input from the frame whose motion is to be protected. We have also seen that such a protection scheme would also prevent detection of pieces of dirt whose areas are comparable to the spatial filter aperture, because they make a significant contribution to the motion trend signal. The size of detectable dirt particles is limited to less than the motion detector spatial filter aperture, so that energy in the frame difference signal caused by dirt is lost in the filter, and does not therefore contribute significantly to the motion trend signal.

Here again, the spatial filter aperture size could be increased to allow detection of larger dirt, but in addition to the appearance of a halo of undetected dirt, the motion detector then becomes blind to slow movement of objects smaller than the filter aperture, some of which will occasionally satisfy the dirt signature and be misinterpreted. An enlarged spatial filter aperture increases the probability of this misinterpretation.

The spatial filter aperture for a centre frame motion detector is therefore a compromise between acceptable motion protection, and adequate dirt detection. An approximately square aperture of 11 field lines by 21 elements ($\approx\frac{1}{34}$ of a picture width) has been found to be a good compromise for 16 mm and 35 mm film, allowing detection of the majority of typical dirt. The remaining motion impairments caused by misinterpretation of slowly moving small detail are reduced to an acceptable level by a "Protection by previous flags" scheme which will be described in Section 3.2.5.

Figure 10 shows one form of centre frame motion detector, whose input is either the difference between centre and previous or centre and future reference frames whichever is the greater. This input function gives better motion protection for a given gain (implemented in the transfer function), than the difference between the centre frame and an equal weighted interpolation between future and past reference frames. Also, the circuits of Fig. 6 and Fig. 10 have the same noise penalty, which simplifies the compensation for film grain which will be mentioned in the following section.

To summarise, this motion detector allows detection of all but the largest pieces of dirt found on typical 16 and 35 mm film whilst at the same time, protecting all motions of objects larger or comparable to its spatial filter aperture, including rapid and erratic motion. It occasionally fails to protect motion of small objects against a plain background, such as rivets in metalwork, causing them to blink out of existence in the concealed picture. This failing can be reduced to acceptable proportions by a protection scheme which will be described in Section 3.2.5.

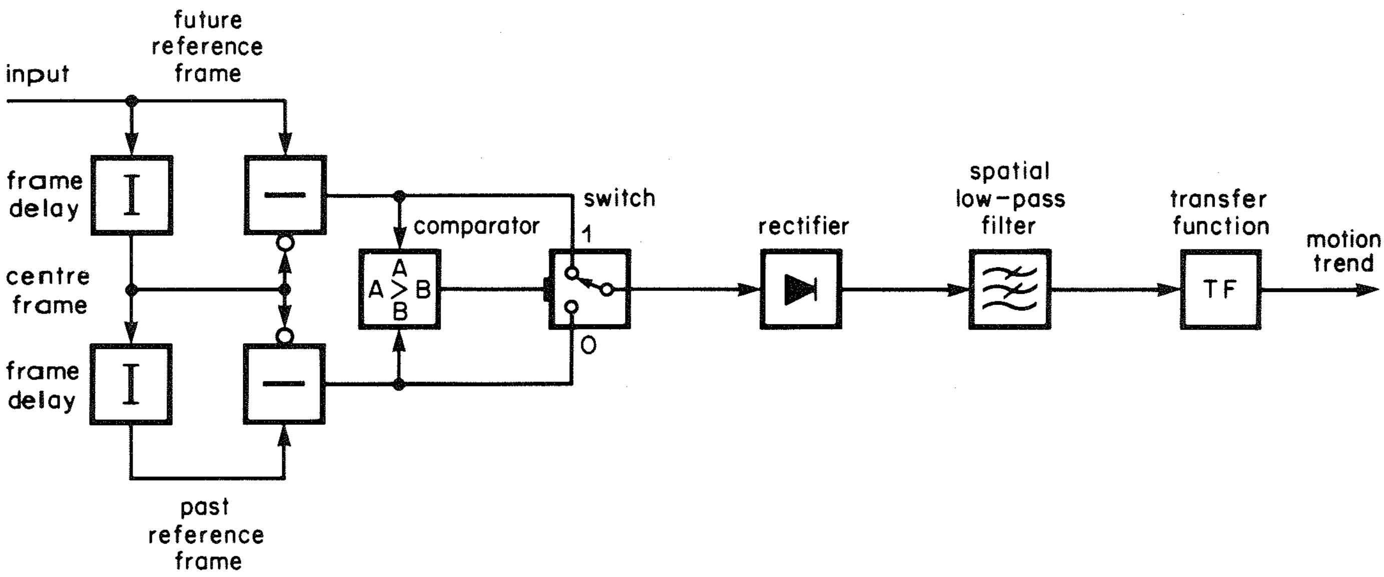


Fig. 10 – Centre frame motion detector.

Such a motion detector could be used in a film dirt concealer for typically dirty film, with motion protection reliable enough not to require previewing of source material.

3.2.4. Combined adaptive motion detector

The centre frame motion detector described above allows effective dirt concealment for typically dirty film. If, however, the film is exceptionally dirty or contains particularly large pieces of dirt, its performance is inadequate.

Under these conditions, an adaptive combination of centre and reference frame motion detectors offers better performance. Reference frame motion detection is used when the scene contains little motion, allowing detection of all sizes of dirt, reverting to centre frame motion detection as the motion content increases.

It is not possible at present, to derive an adaptive control signal as a function of picture content that will guarantee to detect all instances when reference frame motion detection will fail. The adaptive control system chosen for the dirt concealer therefore, offers a set of compromises between motion impairment and dirt concealment effectiveness ranging from the relatively foolproof centre frame motion detector, to an adaptive system capable of excellent dirt concealment at the expense of occasional motion impairment. The choice of a suitable compromise becomes a subjective decision on the part of an operator when a film is being previewed.

The motion detector mode, centre or reference frame, is controlled by a secondary "Global motion detector"[1] and is switched on a field by field basis. Motion of relatively large objects within a scene causes a global motion indication and a corresponding switch to centre frame motion detection.

Figure 11 shows the "global motion detector" which works on the principle of measuring the variation in power in the picture difference signal from line to line. A large variation indicates substantial motion and a small variation generally indicates a static scene.

Referring to Fig. 11, a rectified picture difference signal is integrated for an active line period. The integrator output at the end of a television line corresponds to the mean value of the rectified picture difference signal. A sample and hold circuit stores the integral value so that the integrator can be cleared in preparation for the following line.

The sampled integral consists of two components, the first is proportional to the level of grain in the scene, and the second is a function of the picture difference power caused by motion through the integrated line. It follows that the sampled integral value will vary from line to line. Those in which either no motion exists or motion occurs as a pan across a plain area (e.g. blue sky), will have a small value caused primarily by the rectified grain signal.

Conversely, lines in which significant motion occurs including pans across detailed areas, will have a large integral value corresponding to motion

power and grain level. The contributions from motion and grain are additive, so an approximation to motion power can be made by subtracting the lowest integral value, leaving just the contribution from motion.

Figure 11 shows maximum and minimum detectors following the sample and hold circuit; these retain the largest and smallest integral values respectively, during a film frame. A subtractor takes the difference between maximum and minimum values which is likely, by the end of the frame, to be a reliable approximation to the extent of motion.

The difference and minimum values are stored at the end of the frame for use during the following frame. The minimum value, G, is used to compensate for variations in motion detector sensitivity with grain level (see Appendix 1) and helps to maintain dirt detection sensitivity for very grainy film. The difference signal is compared with three threshold signals, T_1, T_2 and T_3 to provide three control signals P_1, P_2 and P_3, updated once per frame since a new difference signal is stored once per frame. Of the three control signals, P_1 is used to switch between centre and reference frame motion signals as a function of motion content. The remaining two signals, P_2 and P_3 are used to progressively introduce the protection from previous flags system described in the following section.

An operator control varies the threshold values, changing the point at which motion and protection systems are switched, thereby changing the dirt detection effectiveness versus motion protection compromise, according to a subjective assessment of scene content.

Global motion is assessed using signals taken from future reference and centre frames. A result appears at the beginning of the following frame by which time the original source frames have become centre and previous reference frames respectively, so the result is still valid.

3.2.5. Protection by previous flags

This protection scheme largely conceals the effects of misinterpretation of small slowly moving objects against a plain background, when the trend removal signal comes from the centre frame motion detector of Section 3.2.3. Under these conditions, the only contribution to the frame difference signal and therefore the motion trend signal, comes from the moving object itself. If the area of the resulting frame difference signal is sufficiently small compared to the spatial filter aperture (see Appendix 2), the motion trend signal amplitude may not be high enough to prevent misinterpretation of the non-overlapping parts of the image of the object in the centre frame. The subjective effect on a concealed picture is the disappearance of the small object during motion and its reappearance when motion stops.

Rivets in metalwork and coruscation in shaded wind blown trees are common examples of this type of motion.

This protection scheme relies on the close spacing, equal to the motion displacement, of the moving object, on successive frames. If the area immediately surrounding a dirt (or misinterpreted motion) flag, contained similar flags on previous frames, it is more probable that the cause was misinterpreted motion rather than dirt.

There are two parameters in this protection scheme, the optimum search area in previous frames surrounding the "dirt flag" under inspection and the number of past frames to be examined for previous flags. The area in the previous frame, in which dirt flags should be sought is equal to the "expected" displacement of object motion between successive frames. This is a somewhat arbitrary constraint and can only be quantified experimentally, since it is solely governed by picture content. The area to be inspected from older frames should, logically, increase in proportion to the number of separating frames. Although enlarging the inspection area increases the effectiveness of motion protection, it also increases the probability of not concealing valid dirt, especially for very dirty film, so a carefully chosen compromise is required.

Previous flags must be sought in more than one previous frame, because unsteady motion can cause intermittent misinterpretation of slow motion and consequently, a temporary failure of the protection scheme.

A system using flags from the two immediately preceding frames was chosen since it protects unsteady motion with isolated missing flags, without incurring a serious penalty in valid dirt detection, except for extremely dirty film.

A good compromise for previous flag inspection areas was found to be $\frac{1}{44}$ of a picture height by $\frac{1}{32}$ of a picture width for the previous frame and $\frac{1}{34}$ of a picture height by $\frac{1}{16}$ of a picture width for the frame before that. It also proved beneficial to introduce protection progressively, beginning with previous frame only, and progressing to both with increasing global motion, in a similar manner to that used for the adaptive motion detection scheme of section

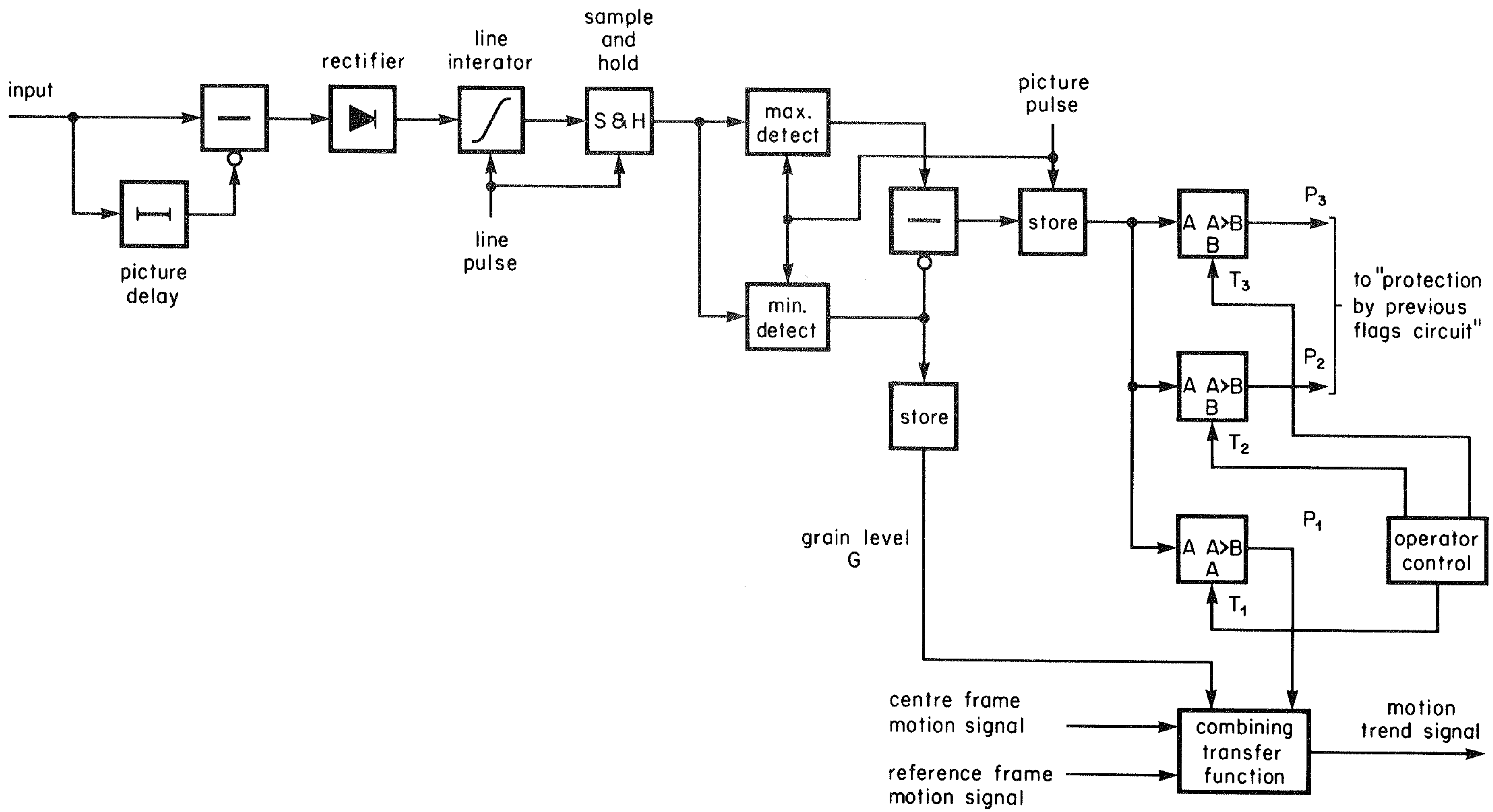


Fig. 11 – "Global motion" detector.

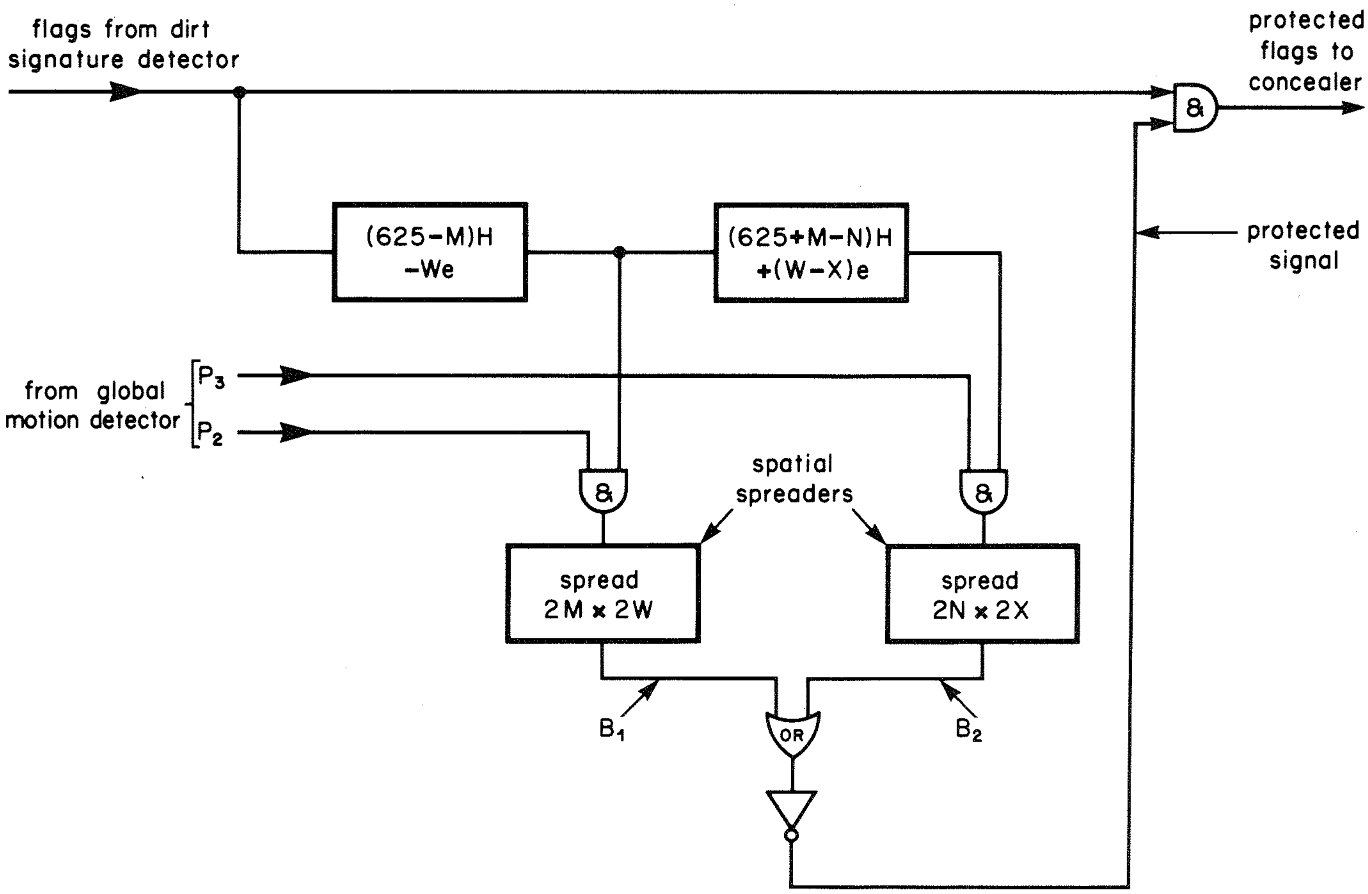


Fig. 12 – Protection by previous flags system.

3.2.4. The global motion thresholds are variable as in the adaptive motion detector, allowing the compromise between protection of small moving detail and effectiveness of dirt detection for very dirty film to be changed for the best subjective effect.

Figure 12 is a block diagram of the "protection by previous flags" system. Flags from the previous two frames are provided by two delays whose lengths are adjusted to compensate for the propagation delays of their respective spatial spread functions. The spatial spreaders generate a two state signal which is valid for a chosen number of lines and elements (a spatial aperture) following an isolated flag stimulus, (see Fig. 13(a)). The compensating delay required to align the box symmetrically about the corresponding pixel in the following frame is a frame delay less half of the spreader spatial aperture.

Figure 13(b) illustrates the protection afforded to a "dirt flag" from the current frame when similar flags exist in the preceding two frames. Since there are three closely spaced flags on successive frames, their most likely cause is slow motion, and the current flag is overridden, thus preserving the image of the moving object. Note that any valid dirt flag falling within the area of either of the spread flags will also be overridden, causing a failure to conceal dirt.

3.3. Unsuccessful motion protection schemes

Several schemes were investigated in the hope of overcoming the motion impairments of the otherwise very promising "Reference Frame" motion detector of section 3.2.2. One of these schemes, the "Centre Frame" motion detector was found to be the most suitable but, for completeness, the unsuccessful attempts are described briefly below. There were four alternative schemes in all, based on processing either the reference frame difference signal or a two level signal derived by amplitude threshold detection of the same signal.

Section 3.2.2, Fig. 9 illustrates the failure of the reference frame motion detector when confronted with rapid motion; that is motion in which the images of a moving object in successive frames do not overlap. The motion trend signal has a large amplitude at the image positions in future and past reference frames but a low value at the centre frame image position. The dirt signature detector therefore, misinterprets the image as dirt with

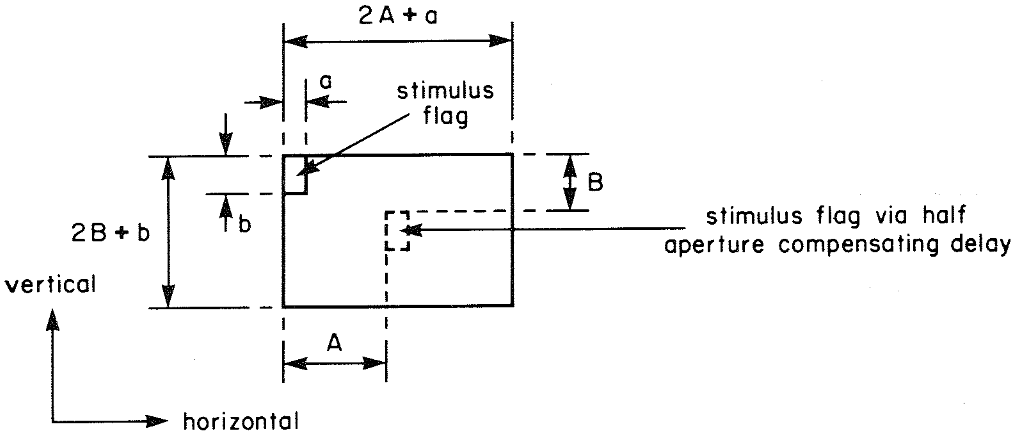


Fig. 13a – Stimulus flag spread spatially with an aperture of 2A × 2B.

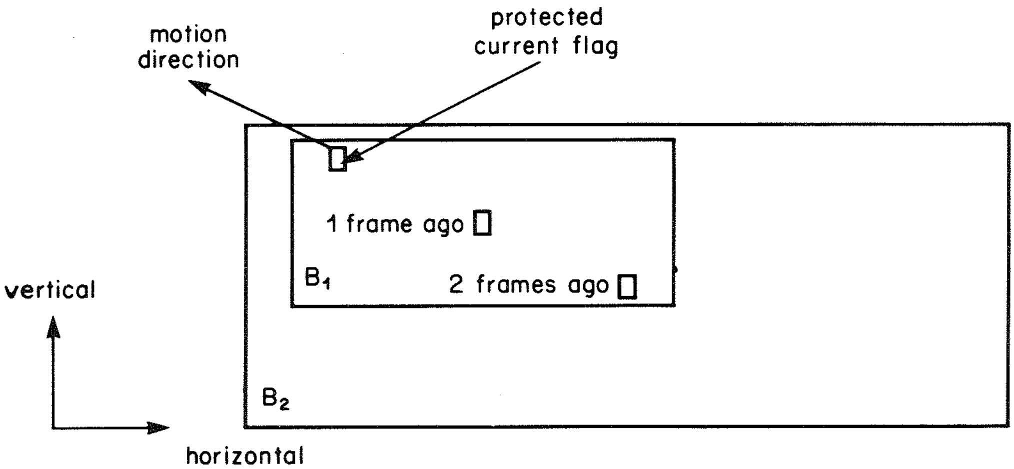


Fig. 13b – Action of the protection circuit.

catastrophic results in the concealed picture. The problem was to find some means to fill the gap between previous and future frame contributions to the motion trend signal and thereby prevent misinterpretations.

3.3.1. Spread threshold

The first scheme derived a two level logical signal by amplitude threshold detection of the rectified reference frame difference signal. The two level signal was expanded horizontally to fill the gap as in Fig. 14(a) and thereby protect the centre frame image by overriding its dirt flag. Expansion must be symmetrical about each thresholded contribution because there is no a priori knowledge of its chronological origin or of the direction of motion. Also, the maximum permissible speed of motion is proportional to the distance of spread.

This scheme was unsatisfactory for two reasons. The first and most significant was the difficulty in finding a satisfactory fixed threshold value, the ideal value being a function of grain level and relative contrast of moving image content. Secondly, the amount of spread required to cover common speeds of rapid motion gave rise to a noticeable halo of unconcealed dirt surrounding moving objects, which would get bigger when the system was expanded to include vertical motions.

The spread threshold could not readily be used to control an adaptive switch between centre and reference frame motion detection because of the threshold problem, also the amount of spread required to cover all possible motions would be so large that its behaviour would be similar to that of the global motion detector.

3.3.2. Large spatial filter

A second approach to filling in the gap between previous and future positions of a moving object was to pass the rectified reference frame difference signal through an equal weighted low pass spatial filter with a very wide aperture. The spatial filter output had a finite value for half of its spatial aperture each side of the input stimulus thereby filling the gap as shown in Fig. 14 (b).

This was tried using a variable horizontal aperture spatial filter with a maximum value of $\frac{1}{6}$ of a

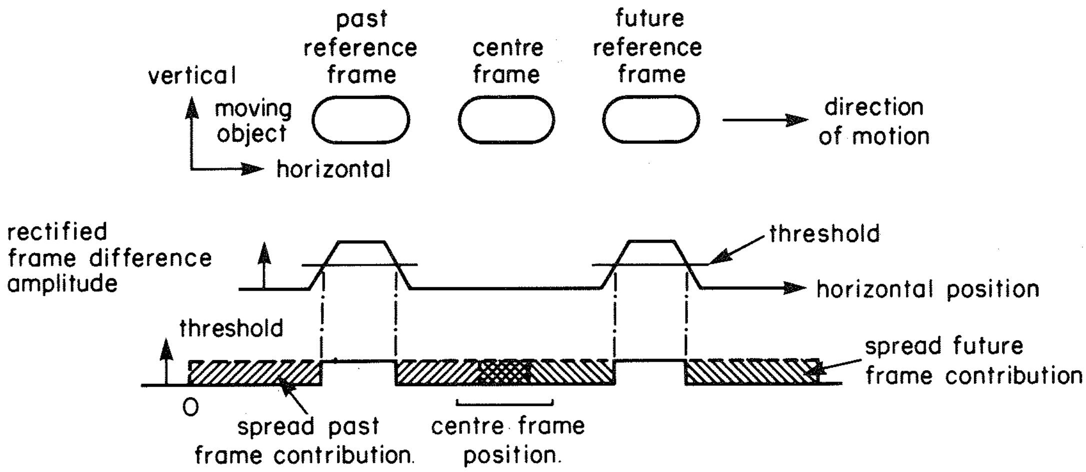


Fig. 14a – Spread threshold motion detector.

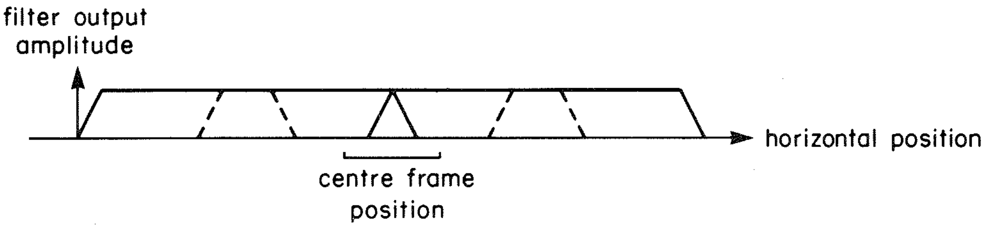


Fig. 14b – Large spatial filter motion detector.

picture width. The vertical aperture was 5 field lines but would need to be increased in proportion to the horizontal aperture in a realistic system, to cope with vertical motion.

When the spatial filter output was used as a motion trend signal, the threshold problems of the "spread threshold" system were replaced by a contradiction over the optimum spatial filter aperture size. If the aperture was large enough to protect typical rapid motion, the motions of objects relatively small in comparison to the aperture (e.g. lamp posts) were not detected. Additionally, an objectionable halo of unconcealed dirt appeared around moving objects.

3.3.3. Line integral

This is an example of a system which works for motion of an idealised plain object against a plain background but fails for motion in real scenes. The unrectified frame difference is integrated along a line as shown in Fig. 15, the integral magnitude is high between the first and second frame difference contributions, effectively filling the gap.

Problems arise however, in realistic scenes, where background detail is concealed or revealed by a moving object or where that object enters or leaves the scene. These events contribute monotonically to the integral, to give a non-zero final value. Although the integral signal can protect against "well behaved" rapid motion, it cannot be used as a motion trend signal because its residual value is too large for most pictures.

3.3.4. Flag correlation

The role of image correlation in predicting positions of moving objects from information in surrounding frames is mentioned in Section 1. Real time image correlation at full resolution, that is using every image sample in the correlation aperture at every possible displacement, requires an impracticably high processing power. Furthermore, correlation can determine a movement vector only for simple translations in which the aspect of a moving object does not change significantly. Real images contain rotations, zooms and a host of similar changes in aspect, so even if it were feasible, correlation could not provide a comprehensive solution to motion problems.

Observations of the "dirt" flags caused by misinterpreted rapid motion showed a succession of approximately equispaced, similarly sized flags at the misinterpreted image positions. The flags were not generally identically shaped, since the causal

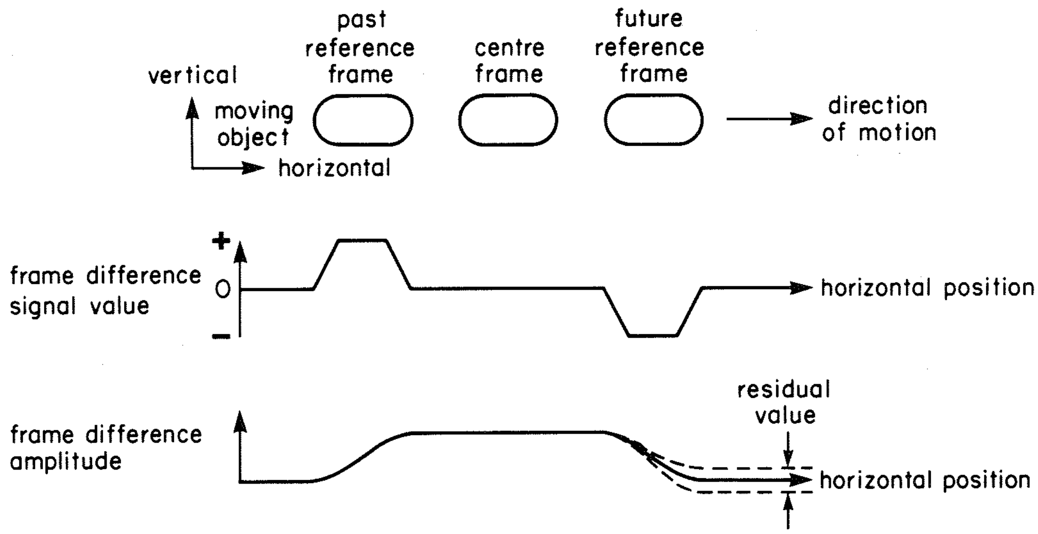


Fig. 15 – Line integral motion detector.

object often changed shape in motion (e.g. a flying bird), and motion often occurred against a background of varying relative contrast.

Nevertheless, it was possible to discern the paths of moving objects from their flag signals, suggesting that a flag correlation scheme might prove successful. If flags taken from two previous frames are correlated to find their relative displacement, a protection signal can be formed in a predicted current frame position and thereby prevent misinterpretation.

An estimate of processing complexity for a real time correlator can be made using the expression:

$$C = B_0 . B_1 . A_H . A_V . D_H . D_V$$

where C = Processing power estimate
B_0 = Number of bits representing first input
B_1 = Number of bits representing second input
A_H = Horizontal aperture in pixels
A_V = Vertical aperture in pixels
D_H = Number of required displacements horizontally
D_V = Number of required displacements vertically

A number of assumptions were made about the nature of the flag signals in order to minimise processing power and thereby make the correlation scheme tractable. An economy of a factor of 64 occurs automatically by correlating two level (one digital bit) flag signals rather than full resolution image signals, which are normally represented by an eight bit signal with 2^8 possible values.

Economies in aperture and displacement contributions can only be made by subsampling the input signals, which in turn requires pre-filtering of input information to guarantee that its bandwidth is less than that which can be supported by the new, lower sampling rate. The information from which flag signals corresponding to motion are derived, is temporally filtered by camera integration and, since the camera shutter angle is normally slightly less than 180°, misinterpreted flags from successive frames can be expected to have an approximately 1:1 mark-space ratio. That is the gap between flags along the motion vector, is approximately equal to the flag dimension along that vector. See Fig. 16. This spatial spreading along the motion vector, as a

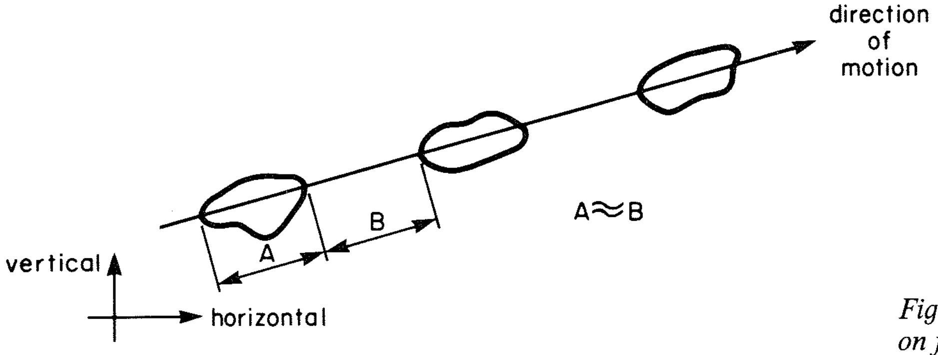


Fig. 16 – Effects of camera integration on flag mark-space ratio.

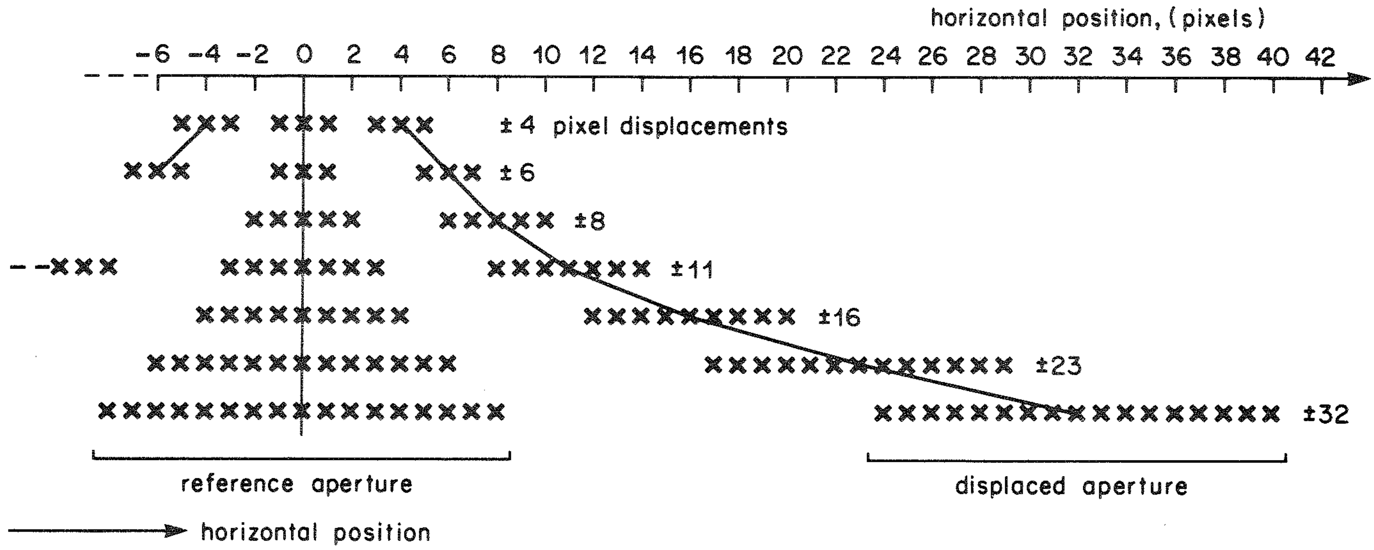


Fig. 17 – Apertures for the flag correlator (single sided displacement).

result of camera integration, amounts to a pre-filter and suggests that either the data sampling rate could be reduced with increasing displacement, or the displacements could be progressively undersampled for increasing values while maintaining the full data sampling rate. The latter approach was chosen for instrumental simplicity.

Correlations for a set of overlapping, sub-sampled horizontal displacements were performed by summing the number of matching logic '1's falling in each aperture – see Fig. 17. Note, also that the aperture sizes are scaled according to displacement. A positive correlation, for any displacement, is declared when its sum exceeds a corresponding threshold. Vertical displacements are undersampled in a similar manner by combining flags from various vertical offsets in a logical "OR" gate and subsequently correlating in the horizontal dimension.

The correlator covers displacements up to ± 32 pixels horizontally and ± 15 field lines vertically, albeit to the limited precision of ± 7 subsampled displacements horizontally and ± 5 vertically. An average of 11.5 pixels of horizontal aperture and 11 pixels of vertical aperture are used compared to 33 and 31 for an equivalent full resolution correlator. Displacement undersampling therefore, gives a further decrease in processing power of $(11.5 \times 11 \times 14 \times 11)/(33 \times 31 \times 64 \times 31) \simeq \frac{1}{104}$. {Each term in brackets is $(A_H \times A_V \times D_H \times D_V)$; there was no correlator for zero horizontal displacement}.

Correlation of flag signals using this scheme rather than full resolution image correlation yields an overall reduction in complexity of approximately 6,600 times. The correlator had a component count of approximately 1,200 integrated circuits, suggesting that real time full resolution image correlation might have a component count somewhere in the region of 10^7, using current technology.

Positive correlations were recombined using a tapped delay network, to reconstruct a protection signal whose dimensions were greater than their causal flags whilst covering their positions in the current frame, assuming steady linear motion.

The correlator worked well for steady linear motion of high contrast objects against a plain background. Under less favourable conditions, however, the flag inputs were less reliable. Background detail and changes of object shape caused flags to disobey the mark-space assumptions of Fig. 16 and occasionally disappear altogether. Also, motion was seldom linear or of constant velocity; and the reduction in aperture for small displacements, combined with one bit input resolution, made the system prone to random correlations in "busy" detailed areas.

The correlator's performance was not good enough to be worth incorporating in the dirt concealer but nevertheless, considering the rather extreme assumptions used in its design, it holds promise for future work.

4. Synthetic dirt modelling

A major problem in assessing the effectiveness of the systems investigated in this project and the

effects of changes in their parameters, was the absence of an analytical measure of dirt detection efficiency. During the early days, assessments were entirely subjective, consisting of capturing three frames from a moving sequence, identifying dirt, and optimising the system parameters for that three frame sample. The process was repeated for several captured sequences and an average setting calculated, which was finally assessed on a continuous moving sequence; a very laborious and error prone procedure.

This problem was overcome to a certain extent using a system which modelled the behaviour of film dirt. A random event generator was developed which generated, at random positions, the density profiles of pieces of dirt whose size distribution was carefully weighted to match the distribution of dirt size found in typical 16mm film. These density profiles were then used to multiply the incoming, relatively clean, film signals in such a way as to model black or printed white dirt.

A one bit field store recorded the positions where synthetic dirt had been inserted. These stored dirt positions were compared with the dirt detector output flag for the appropriate frame and the number of pixels successfully detected, missed and misinterpreted were totalled separately.

The original intention was to use these analytical measurements as error inputs to a closed optimisation loop, to find the set of system parameters giving the best results. Although all of the hardware to implement an optimisation system was available, it was not pursued because, with the aid of the measurement system, manual optimisation was relatively easy and provided a better understanding of the system behaviour.

5. Circuit realisation

5.1. Dirt detection

All data inputs to the dirt detection and concealment circuitry are provided by a 64 K RAM store. Information from four consecutive fields belonging to three consecutive film frames, centred on the frame under inspection, is arranged to appear in order at the read ports R_1 to R_4 of Fig. 18. The output information changes to four fields centred upon the following film frame, once every frame period during dynamic operation, or remains centred upon a single frame when a frozen sequence is required.

Table 1 shows the store addressing sequence and assumes that addresses are subdivided into four blocks of one field capacity indexed A, B, C and D. Each read port column shows the output field along with its index and the write column shows the index under which the current incoming field is written. Note that in the output signal from port 2 an odd output field is out of sequence. This is not significant since this signal is not used for concealment during odd output fields and the detection system operates only during even output fields.

Colour information is removed from the store outputs by subcarrier notch filters before being fed to an array of line delays, configured to provide access to all the points in vertical-temporal space required by the dirt detection circuits. The bulk of the line delays are required to provide access to signals spread over the 11 line vertical aperture of the movement detector spatial filter.

A novel realisation is used for the large aperture spatial filter to minimise hardware requirements.

Table 1 – Store addressing sequence.

Output field no.	Read 1 Field	Read 1 Index	Read 2 Field	Read 2 Index	Read 3 Field	Read 3 Index	Read 4 Field	Read 4 Index	Write Field	Write Index
2 even	1	A	2	B	3	C	4	D	5	B
3 odd	1	A	5	B	3	C	4	D	6	A
4 even	3	C	4	D	5	B	6	A	7	D
5 odd	3	C	7	D	5	B	6	A	8	C
6 even	5	B	6	A	7	D	8	C	1′	A
7 odd	5	B	1′	A	7	D	8	C	2′	B
8 even	7	D	8	C	1′	A	2′	B	3′	C
1′ odd	7	D	3′	C	1′	A	2′	B	4′	D
2′ even	1′	A	2′	B	3′	C	4′	D	5′	B

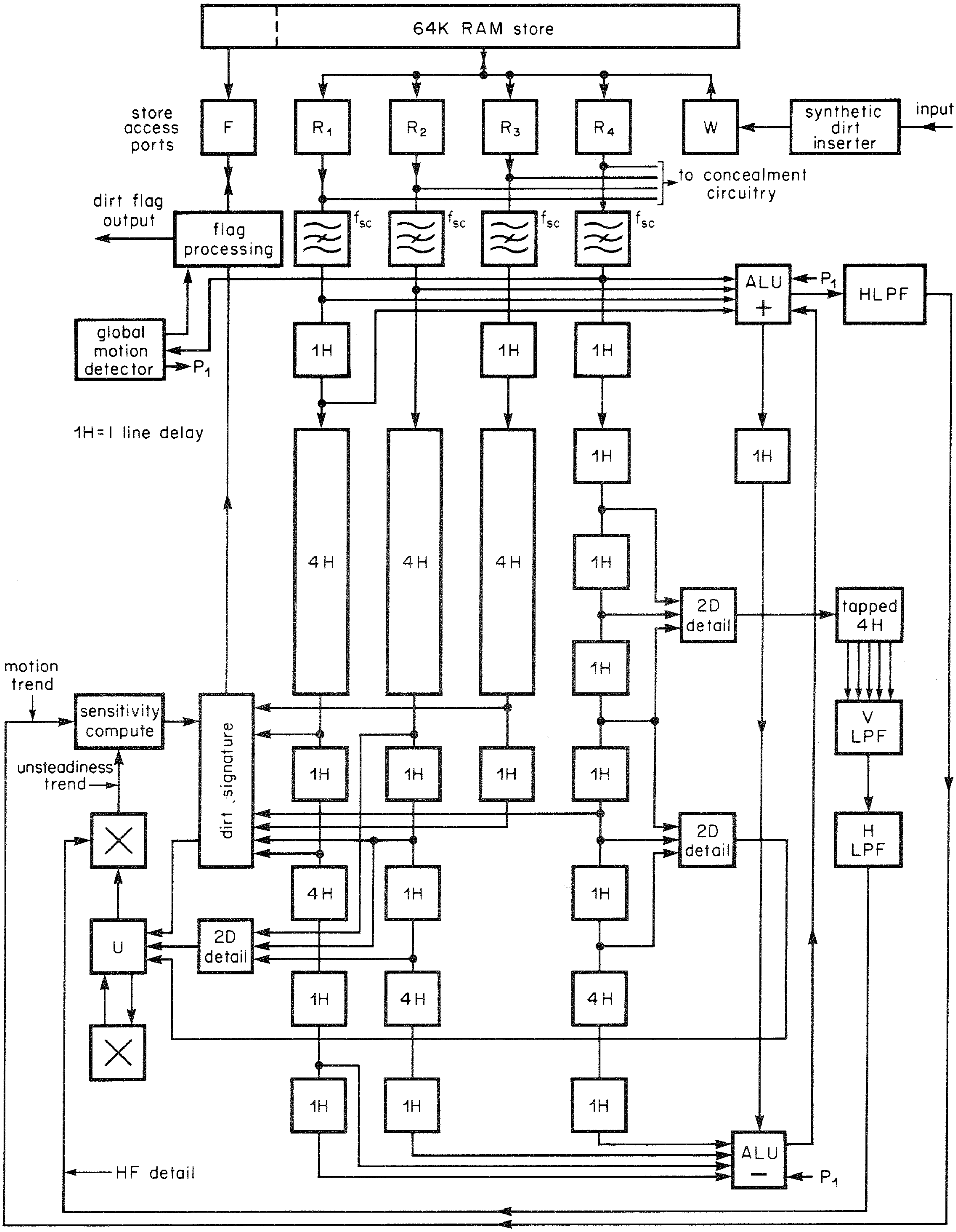


Fig. 18 – Film dirt detection system block diagram.

The filter is realised as separate vertical and horizontal sections each based on the same principle of accumulation and subtraction across a delay equal to the desired aperture. An arithmetic logic unit (A.L.U.) derives either "centre frame" or "reference frame" differences which are rectified and accumulated in separate addresses in a recirculating line delay. Eleven lines later, identical frame differences are subtracted from their corresponding address locations so that each line delay address contains an equal weighted sum of eleven frame differences belonging to a vertical line of eleven pixels. This gives a sin x/x vertical low pass filter with its first zero at 26 c/ph.

The horizontal filter is realised in a similar manner, but using a delay of 21 sample periods to give a sin x/x horizontal filter whose first zero is at 34 c/pw. (giving an approximately square aperture).

The spatial filter output is fed to a "sensitivity calculator" which calculates a movement trend signal from motion and unsteadiness information.

Further tapping points in vertical and temporal space supply inputs to the dirt signature detector and the two detail detectors used for unsteadiness detection. A secondary feed of the rectified difference between centre and future reference frames is supplied by the dirt signature detector and completes the complement of inputs required for unsteadiness detection.

The computed unsteadiness value is used to multiply a spatially filtered detail signal, derived from a third detail detector and spatial low-pass filter combination. This provides an unsteadiness trend contribution in the detailed areas where unsteadiness would otherwise cause misinterpretations.

The combined trend signal is fed to the dirt signature detector where it is used as a threshold beyond which the two opposing frame differences of Fig. 2 must pass for dirt to be detected.

Flag signals are stored in an additional section of the 64 K RAM store. Protection by previous flags as described in section 3.2.5 is performed using flags from previous frames retrieved from this store, to produce a current concealment flag. This processed flag is also stored for retrieval during odd output frames, so that both odd and even images of a piece of dirt can be concealed, even though detection occurs only during the first field of a frame.

5.2. Concealment

The concealment signal is interpolated from stored lines in adjacent frames, one each side of the frame being processed. The interpolation is equally weighted between the two adjacent frames. No attempt is made to find which frame would give the best approximation, because dirt concealment occurs mainly in relatively stationary areas, where inter-frame differences in picture content are small. The improvement in concealment obtained by varying the weights of contributions, would therefore be marginal.

The dirt concealment equipment works on System I composite PAL signals whose colour subcarrier phase repeats over an eight field cycle. Derivation of a concealment signal with a subcarrier phase appropriate to the concealed frame from adjacent frames elsewhere in the eight field cycle requires the use of prediction techniques[5] described below.

Figure 19 shows the positions in vertical-temporal co-ordinates of the information on four successive fields. Also shown are the approximate subcarrier phases relative to the reference point F, which is a line in the first field of the centre frame for which a concealment signal is sought.

Both luminance and chrominance components of the concealment signal are predicted from the previous frame odd field and the future frame even field. In this way only four consecutive fields need be

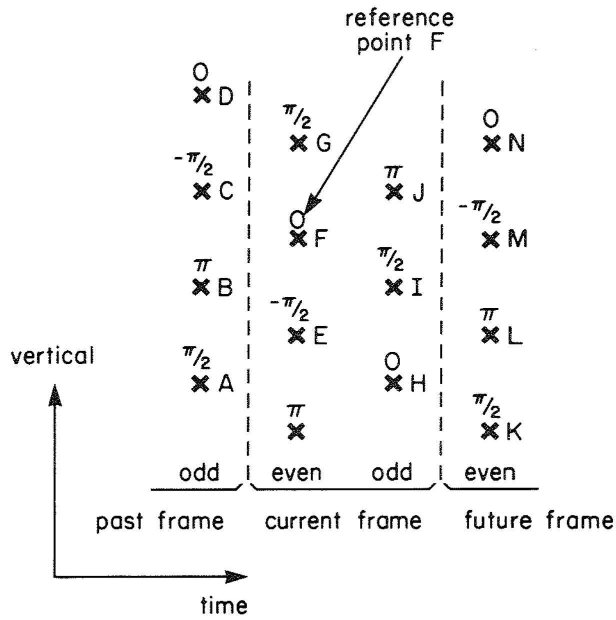


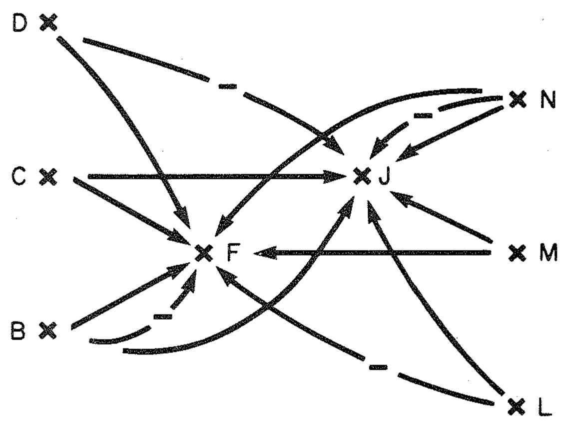
Fig. 19 – Vertical-Temporal positions for concealment.

stored. The luminance component for the even frame is the sum of the spatially coincident line in the future frame and, since there is no spatially coincident line in the past frames, an interpolation from the lines immediately above and below, in the past frame. (See Fig. 20).

A luminance interpolation for the odd field is derived in a similar manner but the single and double contributions are reversed. The resulting prediction is spatially accurate and as temporally accurate as possible, but is vertically low pass filtered by virtue of the vertically displaced contributions.

Chrominance predictions are taken from the nearest lines having the same subcarrier phase and also those having a subcarrier phase difference of π radians, which after signal inversion gives an approximately correct phase. See Fig. 20. The resulting predictions are as temporally correct as possible, but displaced vertically by $\pm\frac{1}{2}$ a picture line. They are also vertically low pass filtered but nevertheless, form an adequate prediction; the 25 Hz vertical hop caused by the vertical displacement is not immediately obvious, being concealed to a large extent by the vertical low-pass filtering.

The luminance prediction is band-stop filtered to remove the remaining incorrect colour information and a matching band-pass filter is used to remove residual luminance components from the chrominance prediction before combining the two



point F $\begin{cases} \text{LUM} = (\text{B} + \text{C} + 2\text{M})/4. \\ \text{CHROM} = (\text{D} - \text{B} + \text{N} - \text{L})/4. \end{cases}$

point J $\begin{cases} \text{LUM} = (2\text{C} + \text{M} + \text{N})/4. \\ \text{CHROM} = (\text{B} - \text{D} + \text{L} - \text{N})/4. \end{cases}$

straight lines ≡ LUMINACE

curved lines ≡ CHROMINANCE

Fig. 20 – Contributions to luminance and chrominance predictions.

to form a PAL coded prediction. A single band-pass filter is used for both luminance and chrominance filters; the band-stop characteristic is formed by inverting the luminance signal before filtering and summing afterwards.

Figure 21 is a block diagram of the predictor and concealment switch which together form the concealment system. Luminance and chrominance

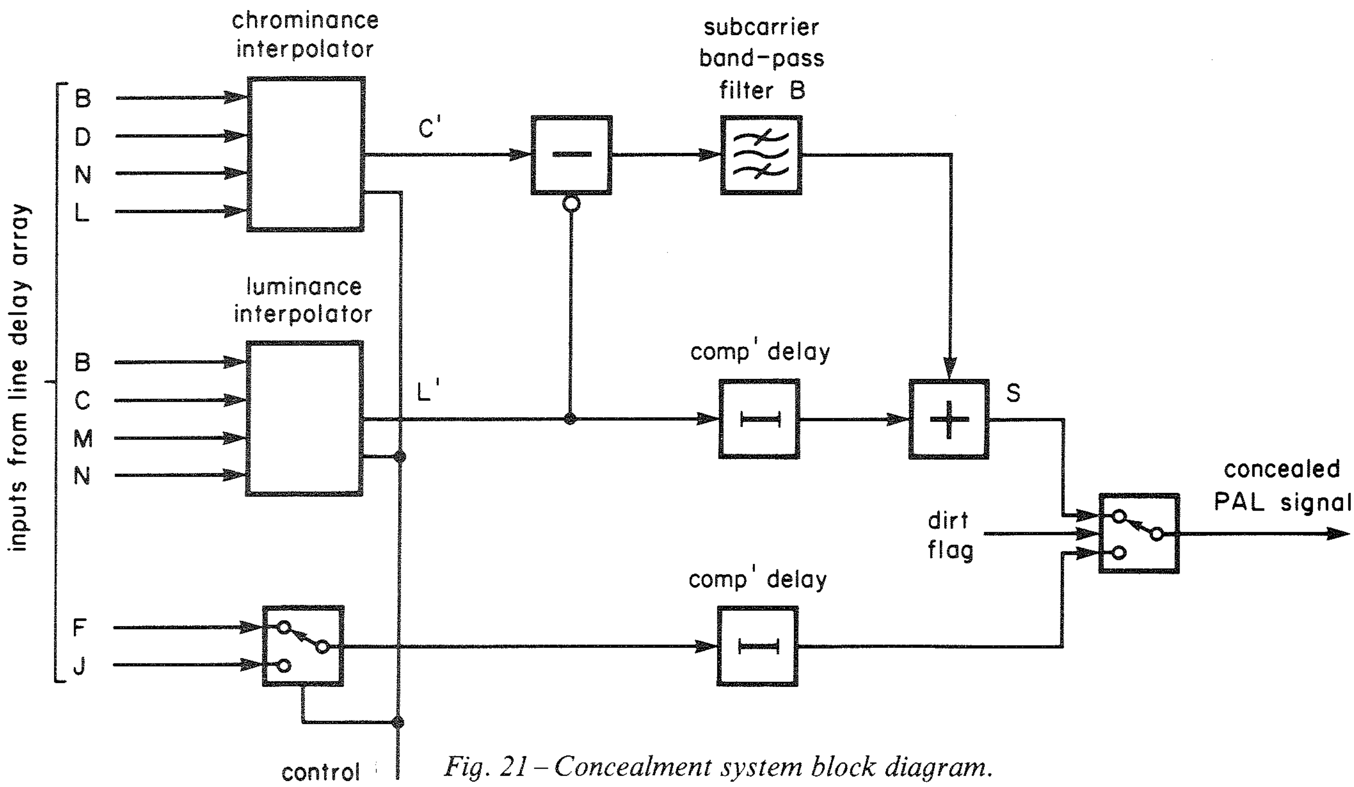


Fig. 21 – Concealment system block diagram.

contributions, L' and C' are formed by separate interpolations to the rules of Fig. 20, and passed through transfer functions of $(1 - B)$ and B respectively. The transfer function B is a bandpass filter centred on the colour subcarrier frequency, so $1 - B$ is a complementary band-stop filter. The overall predictor transfer function is: $S = (1 - B)L' + BC'$ and forms a valid PAL signal. A switch controlled by the "conceal flag" selects between the direct and predicted paths.

6. Conclusions

An electronic film dirt detection and concealment system capable of detecting and concealing dirt in the PAL coded output of a conventional telecine has been described. Its optimum use relies upon manual setting of a compromise between dirt detection effectiveness and motion protection. The compromise is adjustable between detection of dirt with sizes less than 8 pixels in diameter with negligible motion impairments, to detection of dirt of all sizes with occasional motion impairment during certain types of motion.

The equipment may be safely used in its "small dirt only" mode without previewing, when it will conceal the majority of dirt in typical 16 mm film. The remaining large dirt can be concealed by adjusting the motion protection compromise manually, but source material should then be pre-viewed, to check for motion of the type likely to cause misinterpretation. Examples of such motion are birds flying against a uniform or stationary background or small objects such a hailstones moving at medium speeds against a stationary background.

If the large dirt modes are required, the dirt concealer might best be used in the course of a telecine to videotape transfer or during an editing session for film inserts into a studio production.

7. References

1. DREWERY, J.O., STOREY, R., *and* TANTON, N.E., 1984. Video Noise Reduction. BBC Research Department Report No. 1984/7.

2. CHILDS, I., *and* SANDERS, J.R., 1983. New Capabilities for a Line Array CCD Telecine. SMPTE Journal, 1984, 92, 12 pp. 1294–1301.

3. STOREY, R., 1984, Electronic Film Dirt Detection. UK Patent Specification No. 2 139 039.

4. STOREY, R., 1984. "Teletrack"—A Special Effect. BBC Research Department Report No. 1984/10.

5. DREWERY, J.O., *and* STOREY, R., 1979. 8-Field PAL. UK Patent Specification No. 2023366.

APPENDIX 1

Variation of M.D. baseline with noise level

The output of a classical motion detector[1] comprises two components, one is a function of motion and the other is a basal level caused by film grain and varies according to the distribution of grain with scene brightness. There is very little transmission through the film at black and very little emulsion to prevent transmission in white areas, so film grain has a low amplitude at the extremes of transmission and is a maximum at mid-grey.

The presence of this grain contribution adds to the amplitude of the motion trend signal and, if the overall gain of the motion detector for random grain is greater than unity, reduces the likelihood of detecting dirt in very grainy images. This effect can be reduced in an approximate way by subtracting a signal proportional to the mean level of film grain from the motion detector output, thereby removing its offset due to grain.

A measure of mean grain level (G) is available in the global motion detector described in section 3.2.4. A calibration for G was obtained by adding random noise to a stationary noise free picture, and a corresponding correction signal was derived experimentally. The criterion used for choosing a correction signal was a fixed probability of misinterpreting noise peaks for dirt regardless of noise level (and therefore grain in film pictures). A single correction can be used for both centre and reference frame motion detection modes, since their noise penalties are made equal by appropriate choices of frame difference algorithm and filter gain.

APPENDIX 2

Response of a dirt detector with "centre frame" motion detection to idealised dirt of various sizes.

The behaviour of a film dirt detector using a

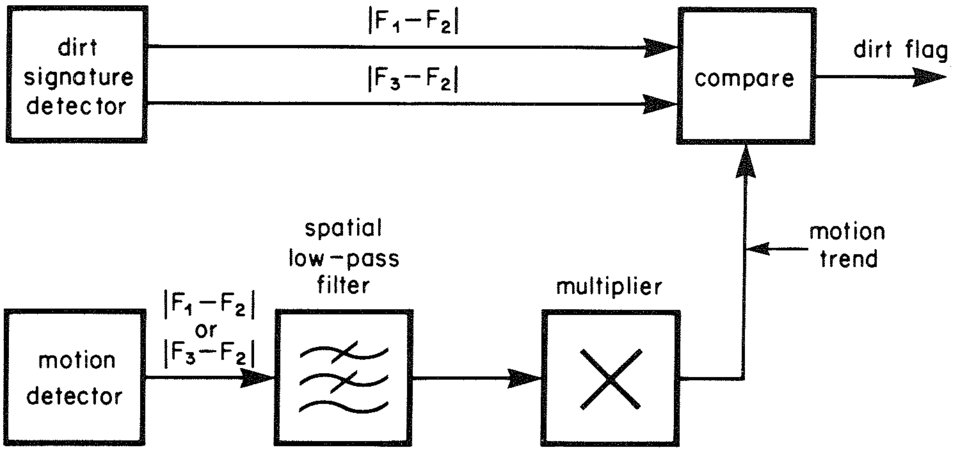


Fig. 22 – Simplified dirt detection circuit.

centre frame motion detector, as a function of dirt size, can be predicted approximately by assuming an idealised uniform density profile for the dirt, and ignoring the effects of the scanning spot aperture and channel bandwidth. This approximation is reasonably good for dirt sizes greater than a few pixels.

A skeleton circuit for the dirt detector is shown in Fig. 22.

The spatial filter output is calculated by convolving its equal weighted aperture with the sample values corresponding to circular dirt of various sizes as the filter aperture scans across the dirt. A uniform density is assumed for the dirt and only the case where it passes through the vertical centre line is considered.

The graph in Fig. 23, shows spatial filter outputs for circular dirt particles with diameters ranging from 3 to 19 pixels. The output values are normalised to the dirt signal amplitude. A multiplier scales the output to form the motion component of the trend removal signal and is set experimentally to give acceptable motion protection. A trend signal of unit amplitude will just prevent detection of valid dirt so the multiplier value also determines the maximum detectable size of dirt.

A multiplier value of 1.125 is found to give acceptable motion protection for the majority of source material and sets the maximum detectable area of dirt to ~50 square pixels, corresponding to circular particles, 8 pixels in diameter. A smaller multiplier value allows detection of larger particles but motion protection was found to be inadequate for values less than 1.125.

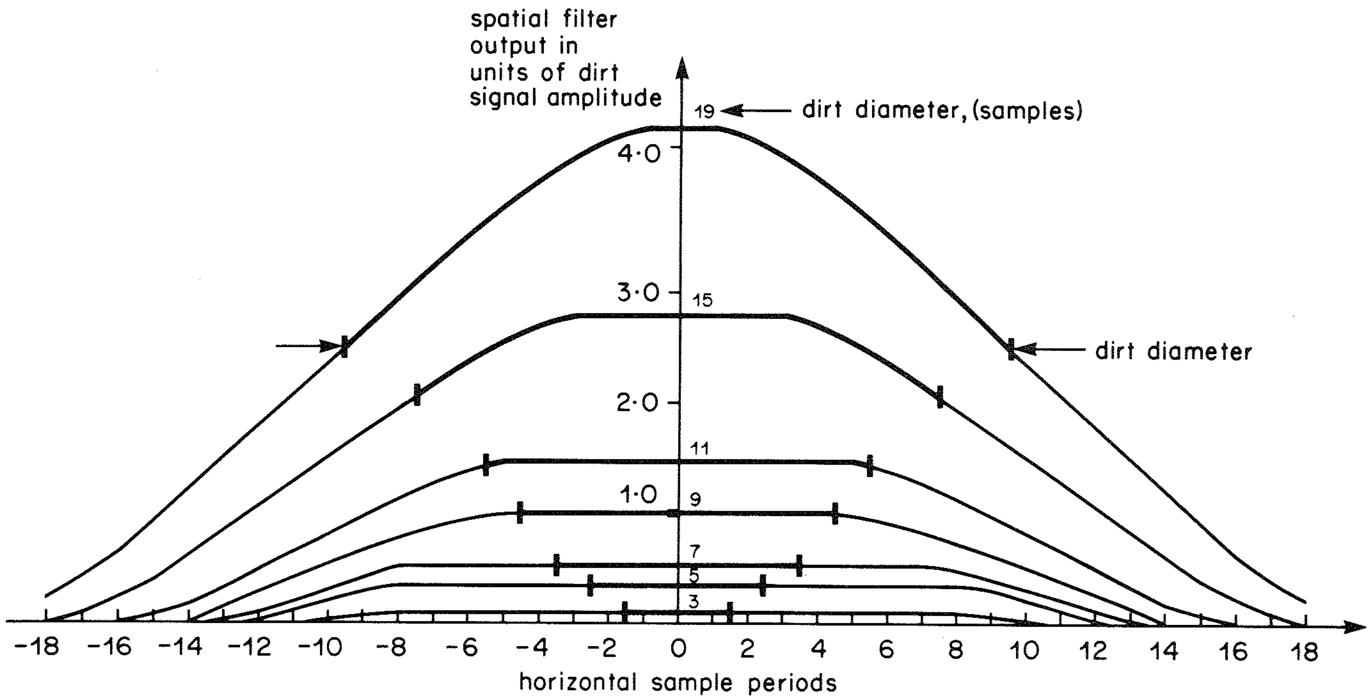


Fig. 23 – Spatial filter output in units of dirt signal amplitude.